(Excerpted from manuscript for Chapter 1 of *RSM Simplified* by Whitcomb and Anderson. See www.statease.com/rsm_simplified.html for information on this book.)

"For a research worker, the unforgotten moments of his [or her] life are those rare ones, which come after years of plodding work, when the veil over nature's secret seems suddenly to lift, and when what was dark and chaotic appears in a clear and beautiful light and pattern."

--Gerty Cori (one of first few women to win the Nobel Prize)

Before we jump down to ground-level details on response surface methods (RSM), let's get a birds-eye view of the lay of the land. First of all, we will assume that you have an interest in this topic from a practical perspective, not academic. A second big assumption is that you've mastered the simpler tools of design of experiments (DOE). (Don't worry, we will do some review in the next few chapters!)

Response Surface Methods (RSM) offer statistical design of experiment (DOE) tools that lead to peak process performance. RSM produces precise maps based on mathematical models. It can put all your responses together via sophisticated optimization approaches, which ultimately lead to the discovery of sweet spots where you meet all specifications at minimal cost.

This answers the question What's in it for me? Now let's see how RSM fits into the overall framework of DOE and provide some historical background.

Strategy of experimentation: The role for RSM

The development of response surface methods began with the publication of a landmark article by Box and Wilson (1951) entitled "On the Experimental Attainment of Optimum Conditions." In a retrospective on events leading up to this paper, Box (2000) recalled observing process improvement teams in the United Kingdom at Imperial Chemical Industries in the late 1940s. He and Wilson realized that, as a practical matter, statistical plans for experimentation must be very flexible and allow for a series of iterations.

Box and other industrial statisticians, notably Hunter (1958-59) continued to hone the strategy of experimentation to the point where it became standard practice in chemical and other process industries in the UK and elsewhere. In the United States, Du Pont took the lead in making effective use of the tools of DOE, including RSM. Via their Management and Technology Center (sadly, now-defunct), they took an in-house workshop called "Strategy of Experimentation" public and, over the last quarter of the 20th century, trained legions of engineers, scientists, and quality professionals on these statistical methods for experimentation.

This now-proven strategy of experimentation, illustrated in Figure 1-1, begins with standard two-level fractional factorial design, mathematically described as "2^{k-p}" (Box, Hunter, 1961), which provides a screening tool. During this phase experimenters seek to discover the vital few factors that create statistically significant effects of practical importance for the goal of process improvement. To save time at this early stage where a number (k) of unknown factors must be quickly screened, the strategy calls for use of relatively low-resolution ("Res") fractions (p).

A QUICK PRIMER ON NOTATION AND TERMINOLOGY FOR STANDARD SCREENING DESIGNS

Two-level factorial design of experiments (DOE) work very well as screening tools. If performed properly, they can reveal the vital few factors that significantly affect your process. To save on costly runs, experimenters often perform only a fraction of all the possible combinations. There are many varieties of fractional two-level designs, such as Taguchi or Plackett-Burman, but we will restrict our discussion to the standard ones that statisticians symbolize as "2^{k-p}", where k refers to the number of factors and minus p is the fraction. Regardless of how you do it, cutting out runs reduces the ability of the design to resolve all possible effects, specifically the higher-order interactions. Minimal-run designs, such as seven factors in eight runs (2^{7-4}) – a 1/16th (2^{-4}) fraction, can only estimate main effects. Statisticians label these low-quality designs as "resolution III" to indicate that main effects will be aliased with two-factor interactions. Resolution III designs can produce significant improvements, but it's like kicking your PC (or slapping the monitor) to make it work: You won't discover what really caused the failure.

To help you grasp the concept of resolution, think of main effects as 1 factor and add this to the number of factors it will be aliased with. In resolution III it's a 1-to-2 relation, which adds to 3. Resolution IV indicates a 1-to-3 aliasing (1+3=4). A resolution V design aliases main effects only with four-factors (1+4=5).

Because of their ability to more clearly reveal main effects, resolution IV designs work much better than resolution III for screening purposes, but they still offer a large savings in experimental run. For example, let's say that you want to screen 10 process factors (k = 10). A full two-level factorial requires 2^{10} (2^k) combinations – way too many (1024!) for a practical experiment. However, the catalog of standard two-level designs offers a 1/32nd fraction that's resolution IV, which will produce fairly clear estimates of main effects. To most-efficiently describe this option mathematically, convert the fraction to 2^p (p = 5) scientific notation: 2^{-5} (= $1/2^5$ = 1/(2x2x2x2x2) = 1/32). This yields 2^{10-5} (2^{k-p}) and by simple arithmetic in the exponent (10-5): 2^5 runs. Now we do the final calculation: 2x2x2x2x2 equals 32 runs in the res IV fraction (versus 1024 in the full-factorial).

PS. A new, more efficient type of fractional two-level factorial screening design has recently been developed (Anderson, Whitcomb, 2004). These designs are referred to as "Min Res IV" because they require a minimal number of factor combinations (runs) to resolve main effects from two-factor interactions (resolution IV). They compare favorably to the classical alternatives on the basis of required experimental runs. For example, 10 factors can be screened in only 20 runs via the Min Res IV whereas the standard (2^{k-p}) resolution IV design, a 1/32nd (2^{-5}) fraction, requires 32 runs.

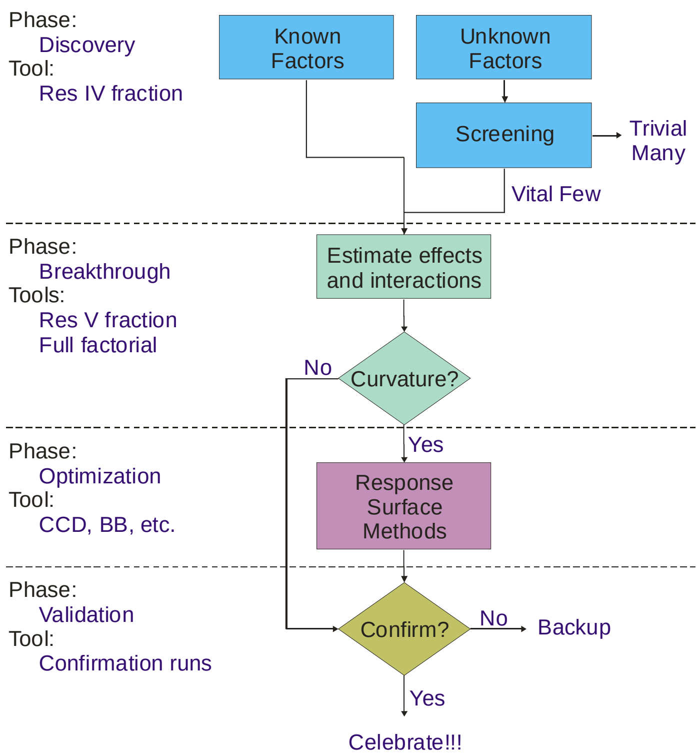

Figure 1-1: Strategy of experimentation

After throwing the trivial many factors off to the side (preferably by holding them fixed or blocking them out), the experimental program should enter the breakthrough phase where interactions become evident. This requires higher-resolution, or possibly full, two-level factorial designs. By definition, traditional one-factor-at-a-time (OFAT) approaches will never uncover interactions of factors that often prove to be the key to success, so practitioners of statistical DOE often generate a huge return-on-investment (ROI) at this breakthrough phase.

As noted in the preface, 80 percent (or more) of all that can be gained in yield and quality from the process might be accomplished at this point, despite having invested only 20 percent of the overall experimental effort. However, high-tech industries facing severe competition cannot stop here. If curvature is detected in their systems, they must optimize their processes for the remaining 20 percent to be gained. As indicated in the flowchart on Figure 1-1, this is the point where response surface methods (RSM) come into play. The typical tools used for RSM, which are detailed later in this book, are the central composite design (CCD) and Box-Behnken design (BBD).

INTERACTIONS – THE HIDDEN GOLD: WHERE TO DIG THE HOLES? (A TRUE CONFESSION.)

The dogma for good strategy of experimentation states that screening designs should *not* include factors known to be active in the system. My co-author Pat, who seldom strays from standard statistical lines, kept saying this to students of Stat-Ease workshops, but I secretly thought he was out of his mind to take this approach. It seemed to me that it was like digging for gold in an area known to contain ore but deliberately doing so in an area far from the mother lode. Finally I confronted Pat about this and asked him to sketch out the flowchart shown above. Then it all made sense to me. It turns out that Pat neglected to mention that he planned to come back to the known factors and combine them with the vital few discovered by screening ones previously not known to have an impact on the system.

- MARK

"It was so dark, he [Stanley Yelnats] couldn't see the end of his shovel. For all he knew he could be digging up gold and diamonds instead of dirt. He brought each shovelful close to his face to see if anything was there, before dumping it out of the hole."

--Holes the Newberry Award-winning children's book by Louis Sachar (1998), made into a popular movie by Walt Disney (2003).